

A Deterministic Algorithm for Maximizing Submodular Functions

Shahar Dobzinski

Ami Mor

July 28, 2015

Abstract

The problem of maximizing a non-negative submodular function was introduced by Feige, Mirrokni, and Vondrak [FOCS'07] who provided a deterministic local-search based algorithm that guarantees an approximation ratio of $\frac{1}{3}$, as well as a randomized $\frac{2}{5}$ -approximation algorithm. An extensive line of research followed and various algorithms with improving approximation ratios were developed, all of them are randomized. Finally, Buchbinder et al. [FOCS'12] presented a randomized $\frac{1}{2}$ -approximation algorithm, which is the best possible.

This paper gives the first deterministic algorithm for maximizing a non-negative submodular function that achieves an approximation ratio better than $\frac{1}{3}$. The approximation ratio of our algorithm is $\frac{2}{5}$. Our algorithm is based on recursive composition of solutions obtained by the local search algorithm of Feige et al. We show that the $\frac{2}{5}$ approximation ratio can be guaranteed when the recursion depth is 2, and leave open the question of whether the approximation ratio improves as the recursion depth increases.

1 Introduction

Let $f : 2^M \rightarrow \mathbb{R}$ be a set function on a ground set M of elements ($|M| = m$). We say that f is *submodular* if it exhibits decreasing marginal utilities. In other words, For every two sets S and T such that $S \subseteq T \subseteq M$ and element $j \in M$:

$$f(S \cup \{j\}) - f(S) \geq f(T \cup \{j\}) - f(T)$$

An equivalent definition is that for every two subsets S and T it holds that:

$$f(S) + f(T) \geq f(S \cup T) + f(S \cap T)$$

It is common to assume that f is non-negative: that is, for every $S \subseteq M$, $f(S) \geq 0$.

Submodular functions arise in many combinatorial optimization scenarios. For example, given a graph $G = (V, E)$, let the cut function of $G : 2^V \rightarrow \mathbb{R}$ be the function that assigns every subset of the vertices S the size of the cut $(S, V - S)$. Observe that a cut function is submodular. Similarly, the objective function in classic problems such as SET-COVER is also submodular. The decreasing marginal utilities property also makes submodular functions a popular object of study in various economic settings, such as combinatorial auctions and influence maximization in social networks (see, e.g., [8] and [7], respectively).

The formal study of approximation algorithms¹ for submodular optimization problems has started back in 1978: Nemhauser, Wolsey and Fisher [9] showed that for monotone and non-negative submodular functions, a natural greedy algorithm gives a $(1 - 1/e)$ -approximation to *monotone* submodular maximization under a cardinality constraint.

¹In general representing a submodular function f takes space that is exponential in m . Therefore, we assume that f is accessed via value queries: given S , what is $f(S)$? All the algorithms we mention make *poly*(m) value queries and run in time *poly*(m).

However, the objective function of problems like $MAX-CUT$ is submodular but not necessarily monotone so the above algorithm does not guarantee any bounded approximation ratio. To handle this lacuna, Feige et al. [3] initiated the study of algorithms for maximizing non-negative submodular functions. They developed a local search algorithm that finds a set whose value is at least $\frac{1}{3}$ of the value of the maximum-value set. In addition, they presented a randomized $\frac{2}{5}$ -approximation. Their algorithms are complemented by the following hardness result: any algorithm that obtains a better approximation ratio than $\frac{1}{2}$, must make exponentially (in m) many queries. This hardness result was later extended [2]: unless $NP = RP$, no $\frac{1}{2} + \epsilon$ polynomial time approximation algorithm exists.

Subsequent work led to improvements in the approximation ratio. Oveis Gharan and Vondrak showed a randomized 0.41-approximation algorithm [6]. Feldman et al. [5] improved this to 0.42, again using a randomized algorithm. Finally, Buchbinder, Feldman, Naor and Schwartz [1] presented a randomized algorithm that achieves an approximation ratio of $\frac{1}{2}$, which as mentioned above, is the best possible.

However, there was no progress at all regarding the approximation ratio achievable by deterministic algorithms, and prior to our work no deterministic algorithm was known to obtain a ratio better than $\frac{1}{3}$. Indeed, this open question was recently mentioned, e.g., by [4]. Our paper makes progress on this front:

Theorem 1.1. *For every $\epsilon > 0$ there is a deterministic algorithm that provides a $(\frac{2}{5} - \epsilon)$ -approximation to unconstrained submodular maximization and makes $O(\frac{1}{\epsilon} \cdot m \cdot \log m)$ value queries.*

Our algorithm can be seen as a recursive composition of solutions obtained by the local search algorithm of [3]. Here is a sketch of our algorithm (see the technical sections for a more accurate description):

1. Find a local maximum² S .
2. Use an approximation algorithm to obtain T_1 with high value when f is restricted to S^c .³
3. Use an approximation algorithm to obtain $T_2 \subseteq S$ such that $f(S^c \cup T_2)$ is high.
4. Return $\arg \max \{S, S^c, T_1 \cup T_2\}$.

For the base case we use in Steps 2 and 3 the local search algorithm of [7]. After obtaining an algorithm with an improved approximation ratio, this new algorithm can be used recursively in Steps 2 and 3. We will show that when the recursion depth is 2 the approximation ratio is $\frac{2}{5}$. The approximation ratio of the algorithm should in principle improve as the recursion depth increases, but we do not know how to analyze it. This is the main open question that this paper leaves.

2 A $\frac{2}{5}$ -Approximation Algorithm

In this section we prove that there exists a $\frac{2}{5}$ -approximation algorithm to the problem of maximizing a non-negative submodular function. To gain some intuition about our algorithm we will see how one can get a better than $\frac{1}{3}$ -approximation (Subsection 2.1). We will then give a formal description and analysis of our algorithm (Subsection 2.2 onwards).

2.1 Warm-Up: Improving over the $\frac{1}{3}$ -Approximation Factor

We now show that we can get an approximation ratio better than $\frac{1}{3}$ approximation. Let S be a local maximum S . Feige et al. [3] show that $2f(S) + f(S^c) \geq f(OPT)$, which implies that either S or S^c provides a

² S is a local maximum if $f(S) \geq f(S \cup T)$ and $f(S) \geq f(S - T)$ for all T .

³We use the notation S^c to denote the complement set of S .

$\frac{1}{3}$ approximation (we let OPT denote the optimal solution). In fact, a slightly stronger inequality holds: $2f(S) + f(S^c) \geq f(OPT) + f(M) + f(\emptyset)$.

Consider a scenario in which the local-maximum algorithm returns only a $\frac{1}{3}$ approximation. In that case we have $f(S) = \frac{f(OPT)}{3}$ and $f(S^c) = \frac{f(OPT)}{3}$. We will show that two other sets, $S^c \cup OPT$ and $S^c \cap OPT$ have high value. For the first set, observe that (using submodularity for the first inequality, S being a local maximum for the second, and $f(S) = \frac{f(OPT)}{3}$ for the third one):

$$\begin{aligned} f(S^c \cup OPT) &\geq f(OPT) + f(M) - f(OPT \cup S) \\ &\geq f(OPT) + f(M) - f(S) \\ &\geq \frac{2f(OPT)}{3} \end{aligned} \tag{1}$$

Similarly $f(S^c \cap OPT) \geq \frac{2f(OPT)}{3}$. In other words:

1. There is a subset of S^c with high value. In particular, the value is at least $2f(S^c)$.
2. There is a superset of S^c with high value. In particular, the value is at least $2f(S^c)$.

Now we would like to (approximately) find these subsets. Observe that we actually have a (smaller) submodular maximization problem. To see this, let $M_1 = S^c$ and $f_1 : 2^{M_1} \rightarrow \mathbb{R}$ be the submodular function $f_1(T) = f(T)$ for every $T \subseteq M_1$. Observe that $f(S^c) = f(M_1) = \frac{f(OPT)}{3}$. In addition, we have already seen that $f(OPT_{M_1}) \geq \frac{f(OPT)}{3}$, where OPT_{M_1} is a maximum value set of f_1 .

We now run the local search algorithm on the function f_1 and find a local maximum $T_1 \subseteq S^c$. By the properties of a local maximum we are guaranteed that

$$2f(T_1) + f(M_1 - T_1) \geq f(OPT_{M_1}) + f(M_1) + f(\emptyset)$$

Notice that by our bounds for $f(OPT_{M_1})$ and $f(M_1)$ it must be that $f(T_1) \geq \frac{f(OPT)}{3}$ or that $f(M_1 - T_1) \geq \frac{f(OPT)}{3}$. If one of them has higher value we already reach an approximation ratio better than $\frac{1}{3}$. Hence, we assume that $f(T_1) = \frac{f(OPT)}{3}$, then $f(M_1 - T_1) = \frac{f(OPT)}{3}$.

We now repeat a similar argument. Similarly to inequality (1),

$$f((M_1 - T_1) \cup OPT_{M_1}) \geq f(OPT_{M_1}) + f(M_1) - f(T_1) \geq \frac{f(OPT)}{3} \tag{2}$$

Let $M'_1 = T_1$. Let $f'_1 : 2^{M'_1} \rightarrow \mathbb{R}$ be the submodular function where $f'_1(T) = f(T \cup (M_1 - T_1))$ for every $T \subseteq M'_1$. Observe that $f'_1(M'_1) = \frac{f(OPT)}{3}$ and $f'_1(\emptyset) = \frac{f(OPT)}{3}$. Furthermore, by inequality (2) the maximum value of f'_1 is at least $\frac{2f(OPT)}{3}$. Let T'_1 be a local maximum of f'_1 . Using [3] again we have that:

$$\max(f(T'_1), f(M'_1 - T'_1)) \geq \frac{1}{3} (f(M_1 - T_1) + f(M_1) + f(OPT_{M_1} \cup (M_1 - T_1))) \geq \frac{4}{9} \cdot f(OPT)$$

This already establishes that the approximation ratio is better than $\frac{1}{3}$, but we can do even better: find $T_2 \subseteq S$ such that $f(S^c \cup T_2)$ has high value. Combine T_1 and T_2 to sum the improvements, since submodularity implies that $f(T_1 \cup T_2) \geq f(S^c) + (f(T_1) - f(S^c)) + (f(S^c \cup T_2) - f(S^c))$.

2.2 The Algorithm

We now present our main result: an improved deterministic algorithm for maximizing a submodular function. As discussed earlier, our algorithm will be based on repeatedly finding local maxima of different subsets and combining the result. However, we do not know how to find a local maximum with a polynomial number of value queries, thus we have to settle on approximate solutions:

Definition 2.1. Fix a function $f : 2^M \rightarrow \mathbb{R}$. S is a $(1 + \epsilon)$ – local maximum of f if for every $T \subset M$:

1. $(1 + \epsilon) f(S) \geq f(S \cup T)$.
2. $(1 + \epsilon) f(S) \geq f(S \cap T)$.

We will use algorithms from [3] and [1] for our “base case”.

Theorem 2.2 (essentially [1, 3]). *There is an algorithm LMS_ϵ that gets as a input function f , a set of elements M and ϵ and returns a set S that is $(1 + \epsilon)$ – approximate local maximum. In addition, $f(S) \geq \frac{f(OPT)}{3}$, where OPT is the set with the highest value. The algorithm makes $O\left(\frac{1}{\epsilon} \cdot m^2 \cdot \log m\right)$ value queries.*

Feige et al [3] show how to start with a set T and find a set S that is a $(1 + \epsilon)$ -approximate local maximum with $O\left(\frac{1}{\epsilon} \cdot m^2 \cdot \log m\right)$ value queries. Their algorithm guarantees that $f(S) \geq f(T)$ and that either $f(S) \geq \frac{f(OPT)}{3(1+\epsilon)}$ or $f(S^c) \geq \frac{f(OPT)}{3(1+\epsilon)}$. This weaker guarantee suffices to our analysis, but it would be simpler to first obtain a set T with $f(T) \geq \frac{f(OPT)}{3}$ via the deterministic algorithm of [1], and then obtain an approximate local maximum S by running the algorithm of [3] with T as the initial set.

We are now ready to present our recursive algorithm for unconstrained submodular maximization. The input is a submodular function f defined on a set of elements M , and a parameter $nrounds$ that determines the recursion depth. We also assume some fixed accuracy parameter $\epsilon > 0$. The algorithm returns a set that will be proved to provide a high value.

$USM_\epsilon(f, M, nrounds) :$

1. Define f' : $f'(T) = f(T) - \min(f(\emptyset), f(M))$. Let $S = LMS_\epsilon(f', M)$.
2. If $S = M$, $S = \emptyset$, or $nrounds = 0$ then return S .
3. Let $M_1 = S^c$, define f_1 : $f_1(T) = f'(T)$. Let $T_1 = USM_\epsilon(f_1, M_1, nrounds - 1)$.
4. Let $M_2 = S$, define f_2 : $f_2(T) = f'(S^c \cup T)$. Let $T_2 = USM_\epsilon(f_2, M_2, nrounds - 1)$.
5. Return $\arg \max_{R \in \{S, (T_1 \cup T_2), M, \emptyset\}} f(R)$.

Theorem 2.3. *Let $f(USM_\epsilon(f, M, 2)) \geq (\frac{2}{3} - \epsilon) \cdot f(OPT)$. Moreover, the algorithm makes $O\left(\frac{1}{\epsilon} \cdot m^2 \cdot \log m\right)$ value queries in this case.*

2.3 Proof of Theorem 2.3

We first analyze the running time of the algorithm. Let $L_\epsilon(m)$ be the maximum number of queries LMS_ϵ makes on any submodular function that is defined on m elements.

Lemma 2.4. *For every integer $n \geq 0$ the number of queries used by $USM_\epsilon(f, M, n)$ is $O(m \cdot L_\epsilon(m))$. As a corollary, for every integer $n \geq 0$ the number of queries used by $USM_\epsilon(f, M, n)$ is $O\left(\frac{1}{\epsilon} \cdot m^3 \log m\right)$.*

Proof. Denote by $T_\epsilon(k)$ the maximal number of queries USM_ϵ makes on a submodular function with $|M| = k$. In iterations where the algorithm stops at Step 2, the number of queries is at most $L_\epsilon(|M|)$. Else, we have that $|M| > |S| > 0$. The algorithm then recursively solves two subproblems, one on S and the other on S^c . Thus, $T_\epsilon(m)$ can be bounded by:

$$T_\epsilon(m) \leq \max_{d=1..m-1} (T_\epsilon(d) + T_\epsilon(m-d)) + L_\epsilon(m)$$

Solving that recursion by induction on m gives that $T_\epsilon(m) \leq L_\epsilon(m) \cdot m$. □

The next lemma bounds the approximation ratio when the recursion depth is 2 (i.e., $nrounds = 2$). We note that using a larger value of $nrounds$ should in principle yield an improved ratio, but currently we are unable to formally analyze this case. The next subsections are devoted to proving the lemma.

Lemma 2.5. *For every $\epsilon > 0$ and for every submodular function f with $f(M) \geq 0$ and $f(\emptyset) \geq 0$:*

$$f(USM_\epsilon(f, M, 2)) \geq \left(\frac{2}{5} - \epsilon\right) f(OPT)$$

2.4 Proof Overview

We introduce a function α_i which represents a lower bound on the value of the solution that the algorithm outputs after i rounds. Let \mathcal{F} be the set of all submodular functions.

Definition 2.6. $\alpha_i(x_{OPT}, x_0, x_M) = \inf_{f \in \mathcal{F}} \{USM_\epsilon(f, M, i) \mid f(OPT) \geq x_{OPT}, f(\emptyset) \geq x_0, f(M) \geq x_M\}$.

Our result is based on the following lower bounds for α_i . The proofs are in Subsection 2.5.

Claim 2.7. *For every non negative x_{OPT}, x_0, x_M :*

$$\alpha_0(x_{OPT}, x_0, x_M) \geq \max\left(\frac{1}{3}(x_{OPT} + x_0 + x_M), x_M, x_0\right)$$

Claim 2.8 (composition lemma). *Let f be a submodular function, $i > 0$, and S, T_1 and T_2 are the sets defined in the algorithm when we run $USM_\epsilon(f, M, i)$. Then,*

$$\begin{aligned} f(T_1 \cup T_2) &\geq \alpha_{i-1}(f(OPT) + f(\emptyset) - (1 + \epsilon)f(S), f(\emptyset), f(S^c)) + \\ &\quad \alpha_{i-1}(f(OPT) + f(M) - (1 + \epsilon)f(S), f(S^c), f(M)) - f(S^c) \end{aligned}$$

And in particular:

$$f(T_1 \cup T_2) \geq \alpha_{i-1}(f(OPT) + f(M) - (1 + \epsilon)f(S), f(S^c), f(M))$$

Claim 2.9. *For every non negative x_{OPT}, x_M, x_0 we have:*

$$\begin{aligned} \alpha_1(x_{OPT}, 0, x_M) &\geq \max\left(\frac{1 - \epsilon}{3} \cdot x_{OPT} + \frac{1 - \epsilon}{2} \cdot x_M, x_0\right) \\ \alpha_1(x_{OPT}, x_0, 0) &\geq \max\left(\frac{1 - \epsilon}{3} \cdot x_{OPT} + \frac{1 - \epsilon}{2} \cdot x_0, x_M\right) \end{aligned}$$

Before proving these two claims, let us see why they let us derive our main bound on the quality of the solution (note that lemma 2.5 can be directly derived from this claim):

Claim 2.10. *For every non negative x_{OPT}, x_M we have:*

$$\alpha_2(x_{OPT}, 0, 0) \geq \left(\frac{2}{5} - \epsilon\right) x_{OPT}$$

Proof. Let f be a submodular function such that $f(OPT) \geq x_{OPT}$, $f(\emptyset) \geq 0$ and $f(M) \geq 0$. We want to show that:

$$USM_\epsilon(f, M, 2) \geq \left(\frac{2}{5} - \epsilon\right) f(OPT)$$

Let T_1, T_2 be the sets defined by the algorithm when we run $USM_\epsilon(f, M, 2)$. Using the composition lemma we know that:

$$f(T_1 \cup T_2) \geq \alpha_1(f(OPT) - (1 + \epsilon)f(S), 0, f(S^c)) + \alpha_1(f(OPT) - (1 + \epsilon)f(S), f(S^c), 0) - f(S^c)$$

We use claim 2.9 to bound α_1 and get:

$$\begin{aligned} f(T_1 \cup T_2) &\geq 2(1 - \epsilon) \left(\frac{1}{3} [f(OPT) - (1 + \epsilon)f(S)] + \frac{1}{2}f(S^c) \right) - f(S^c) \\ &= \frac{2}{3}(1 - \epsilon)f(OPT) - \frac{2}{3}(1 - \epsilon^2)f(S) - \epsilon f(S^c) \\ &\geq \frac{2}{3}(1 - \epsilon)f(OPT) - \frac{2}{3}f(S) - \epsilon f(OPT) \\ &= \frac{2}{3}(f(OPT) - f(S)) - \frac{5}{3}\epsilon f(OPT) \end{aligned}$$

The algorithm compares $f(T_1 \cup T_2)$ to $f(S)$ and returns the set with highest value. Therefore,

$$USM_\epsilon(f, M, 2) \geq \max \left(f(S), \frac{2}{3}f(OPT) - \frac{2}{3}f(S) - \frac{5}{3}\epsilon f(OPT) \right)$$

That is a maximum of two functions, both linear in $f(S)$, one is ascending and one is descending. The minimum of a maximum of two such functions is achieved at their intersection. I.e., when $f(S) = \frac{2}{3} \cdot f(OPT) - \frac{5}{3} \cdot \epsilon \cdot f(OPT)$. Solving this equation gives us:

$$USM_\epsilon(f, M, 2) \geq \left(\frac{2}{5} - \epsilon \right) f(OPT)$$

By the definition of α_2 and since $f(OPT) \geq x_{OPT}$,

$$\alpha_2(x_{OPT}, 0, 0) \geq \left(\frac{2}{5} - \epsilon \right) x_{OPT}$$

□

2.5 Proofs

To finish the proof of our main result all we are left with is proving the bounds for α_0 and α_1 (claims 2.7, 2.9) and the composition lemma (claim 2.8). We start with some helpful observations.

2.5.1 Observations

Let S be a $(1 + \epsilon)$ -local maximum. Therefore:

$$(1 + \epsilon)f(S) \geq f(S \cap OPT) \tag{3}$$

$$(1 + \epsilon)f(S) \geq f(S \cup OPT) \tag{4}$$

Also, submodularity of f implies that

$$f(OPT \cap S^c) + f(OPT \cap S) \geq f(OPT) + f(\emptyset) \tag{5}$$

$$f(OPT \cup S^c) + f(OPT \cup S) \geq f(OPT) + f(M) \tag{6}$$

The next claim appeared in [3] in a slightly different form:

Claim 2.11. *Let S be a $(1 + \epsilon)$ -local maximum, and f a submodular function. Then $2(1 + \epsilon)f(S) + f(S^c) \geq f(OPT) + f(M) + f(\emptyset)$.*

Proof. Since f is submodular,

$$f(S^c) + f(OPT \cup S) \geq f(OPT \cap S^c) + f(M).$$

Inequalities (3), (4), (5) together with the above inequality gives us:

$$\begin{aligned} 2(1 + \epsilon)f(S) + f(S^c) &\geq f(S \cap OPT) + f(S \cup OPT) + f(S^c) \\ &\geq f(S \cup OPT) + f(OPT \cap S^c) + f(M) \\ &\geq f(OPT) + f(\emptyset) + f(M) \end{aligned}$$

□

2.5.2 Proof of Claim 2.7

Let f be a submodular function, and S be the local maximum found by the algorithm. Using theorem 2.2:

$$3f(S) \geq f(OPT) + f(M) + f(\emptyset)$$

Also, note that the algorithm takes the set with highest value among S , S^c , M and \emptyset . Therefore, for every submodular f :

$$USM_\epsilon(f, M, 0) \geq \max\left(f(\emptyset), f(M), \frac{1}{3}(f(OPT) + f(M) + f(\emptyset))\right)$$

Hence, by the definition of α_0 ,

$$\alpha_0 \geq \max\left(x_o, x_M, \frac{1}{3}(x_{OPT} + x_0 + x_M)\right)$$

2.5.3 Proof of Claim 2.8 (composition lemma)

We follow steps 3 and 4 of the algorithm. The algorithm takes the set S^c and attempts to improve it by adding some elements and removing others. Since f is submodular and $T_1 \subseteq S^c$, then $f(T_2|T_1) \geq f(T_2|S^c)$. Therefore:

$$f(T_1 \cup T_2) = f(T_1) + f(T_2|T_1) \geq f(T_1) + f(T_2|S^c) = f_1(T_1) + f_2(T_2) - f(S^c) \quad (7)$$

Next, we give bounds on $f_1(T_1)$ and $f_2(T_2)$:

Claim 2.12. *The above T_1 and T_2 satisfy:*

$$\begin{aligned} f_1(T_1) &\geq \alpha_{i-1}(f(OPT) + f(\emptyset) - f(S), f(\emptyset), f(S^c)) \\ f_2(T_2) &\geq \alpha_{i-1}((f(OPT) + f(M) - f(S), f(S^c), f(M)) \end{aligned}$$

Proof. First, note that by the definition of α_i , it is monotone in the following sense: if $x_{OPT} \geq y_{OPT}$, $x_0 \geq y_0$ and $x_M \geq y_M$, then $\alpha_i(x_{OPT}, x_0, x_M) \geq \alpha_i(y_{OPT}, y_0, y_M)$.

Consider the subproblems in which we attempt to find a maximum of f_1 and f_2 . By the definition of α_{i-1} , we have:

$$f_1(T_1) = USM_\epsilon(f_1, M_1, i-1) \geq \alpha_{i-1} \left(\max_{R \subseteq S^c} (f_1(R)), f_1(\emptyset), f_1(M_1) \right) \quad (8)$$

$$f_2(T_2) = USM_\epsilon(f_2, M_2, i-1) \geq \alpha_{i-1} \left(\max_{R \subseteq S} (f_2(R)), f_2(\emptyset), f_2(M_2) \right) \quad (9)$$

Combining (3) and (5) gives us $f_1(OPT \cap S^c) \geq f(OPT) + f(\emptyset) - (1+\epsilon)f(S)$. Therefore, $\max_{R \subseteq S^c} f_1(R) \geq f_1(OPT \cap S^c) \geq f(OPT) + f(\emptyset) - (1+\epsilon)f(S)$. Note also that $f_1(\emptyset) = f(\emptyset)$ and $f_1(M_1) = f(S^c)$. Applying inequality (8) and using the monotonicity of α_i , we have:

$$\begin{aligned} f_1(T_1) &\geq \alpha_{i-1} \left(\max_{R \subseteq S^c} (f_1(R)), f_1(\emptyset), f_1(M_1) \right) \\ &\geq \alpha_{i-1} (f(OPT) + f(\emptyset) - (1+\epsilon)f(S), f(\emptyset), f(S^c)) \end{aligned}$$

Note that $f_2(OPT \cap S) = f(OPT \cup S^c)$. Combining inequalities (4) and (6) together we get $f_2(OPT \cap S) \geq f(OPT) + f(M) - (1+\epsilon)f(S)$. Therefore, $\max_{R \subseteq S} f_2(R) \geq f_2(OPT \cap S) \geq f(OPT) + f(M) - (1+\epsilon)f(S)$. Note also that $f_2(\emptyset) = f(S^c)$ and $f_2(M_2) = f(M)$. Applying inequality (9) and using the monotonicity of α_i , we have:

$$\begin{aligned} f_2(T_2) &\geq \alpha_{i-1} \left(\max_{R \subseteq S} f_2(R), f_2(\emptyset), f_1(M_2) \right) \\ &\geq \alpha_{i-1} (f(OPT) + f(M) - (1+\epsilon)f(S), f(S^c), f(M)) \end{aligned}$$

□

Back to the proof of claim 2.8, by (7) we have $f(T_1 \cup T_2) \geq f(T_1) + f(T_2) - f(S^c)$. Applying claim 2.12 gives us:

$$\begin{aligned} f(T_1 \cup T_2) &\geq \alpha_{i-1} (f(OPT) + f(\emptyset) - (1+\epsilon)f(S), f(\emptyset), f(S^c)) + \\ &\quad \alpha_{i-1} (f(OPT) + f(M) - (1+\epsilon)f(S), f(S^c), f(M)) - f(S^c) \end{aligned}$$

To prove second part of the composition lemma, we use the following claim:

Claim 2.13. *For every non-negative x_{OPT}, x_M, x_0 and every i we have: $\alpha_i(x_{OPT}, x_0, x_M) \geq x_M$.*

Proof. In Step 2 the algorithm can choose to output M . Therefore, $USM_\epsilon(f, M, i) \geq f(M)$. □

Using claim 2.13 we know that

$$\alpha_{i-1} (f(OPT) + f(\emptyset) - (1+\epsilon)f(S), 0, f(S^c)) \geq f(S^c)$$

Combining it with the first part of the composition lemma gives us:

$$\begin{aligned} f(T_1 \cup T_2) &\geq \alpha_{i-1} (f(OPT) + f(\emptyset) - (1+\epsilon)f(S), f(\emptyset), f(S^c)) + \\ &\quad \alpha_{i-1} (f(OPT) + f(M) - (1+\epsilon)f(S), f(S^c), f(M)) - f(S^c) \\ &\geq \alpha_{i-1} (f(OPT) + f(M) - (1+\epsilon)f(S), f(S^c), f(M)) \end{aligned}$$

2.5.4 Proof of Claim 2.9

Let f be a submodular function such that $f(OPT) \geq x_{OPT}$, $f(\emptyset) \geq 0$ and $f(M) \geq x_M$. We will show that:

$$USM_\epsilon(f, M, 1) \geq \max \left(x_M, \frac{1}{3} (f(OPT) + f(M)) + \frac{1}{6} f(M) \right)$$

Let T_1, T_2 be the sets defined by the algorithm when we run $USM_\epsilon(f, M, 1)$. Using the second part of the composition lemma we know that:

$$f(T_1 \cup T_2) \geq \alpha_0(f(OPT) + f(M) - (1 + \epsilon)f(S), f(S^c), f(M)) \quad (10)$$

Claim 2.7 shows that $\alpha_0(x_{OPT}, x_0, x_M) \geq \frac{1}{3}(x_{OPT} + x_0 + x_M)$, and claim 2.11 gives us: $f(S^c) \geq f(OPT) + f(M) - 2(1 + \epsilon)f(S)$. Therefore,

$$\begin{aligned} & \alpha_0(f(OPT) + f(M) - (1 + \epsilon)f(S), f(S^c), f(M)) \\ & \geq \frac{(f(OPT) + 2f(M) - (1 + \epsilon)f(S) + f(S^c))}{3} \\ & \geq \frac{2f(OPT) + 3f(M) - (1 + \epsilon)3f(S)}{3} \end{aligned} \quad (11)$$

From inequalities (10) and (11) it follows that:

$$f(T_1 \cup T_2) \geq \frac{2f(OPT) + 3f(M) - (1 + \epsilon)3f(S)}{3}$$

The value of the output of the algorithm is at least the maximum of $f(T_1 \cup T_2)$, $f(S)$, \emptyset and M . Hence,

$$USM_\epsilon(f, M, 1) \geq \max \left\{ f(S), \frac{2f(OPT) + 3f(M) - (1 + \epsilon)3f(S)}{3} \right\}$$

This is a maximum of two functions, both linear in $f(S)$, one is ascending and one is descending. The minimum of a maximum of two such functions is obtained when they intersect, i.e., when $f(S) = f(T_1 \cup T_2)$. Therefore $\max(f(S), f(T_1 \cup T_2)) \geq \frac{1}{6+3\epsilon}(2f(OPT) + 3f(M))$. Since $\frac{1}{6+3\epsilon} \geq \frac{1}{6}(1 - \epsilon)$ we have:

$$\max(f(S), f(T_1 \cup T_2)) \geq (1 - \epsilon) \left(\frac{1}{3} \cdot f(OPT) + \frac{1}{2} \cdot f(M) \right)$$

Therefore,

$$USM_\epsilon(f, M, 1) \geq \max \left(f(M), \frac{1 - \epsilon}{3} (f(OPT) + f(M)) + \frac{1 - \epsilon}{6} \cdot f(M) \right)$$

By the definition of α_1 , it follows that:

$$\alpha_1(x_{OPT}, 0, x_M) \geq \max \left(x_M, \frac{1 - \epsilon}{3} (x_{OPT} + x_M) + \frac{1 - \epsilon}{6} \cdot x_M \right)$$

The proof that $\alpha_1(x_{OPT}, x_0, 0) \geq \max(\frac{1 - \epsilon}{3} \cdot x_{OPT} + \frac{1 - \epsilon}{2} \cdot x_0, x_0)$ is similar, with the roles of M and \emptyset switched.

References

- [1] Niv Buchbinder, Moran Feldman, Joseph Naor, and Roy Schwartz. A tight linear time $(1/2)$ -approximation for unconstrained submodular maximization. In *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*, pages 649–658. IEEE, 2012.
- [2] Shahr Dobzinski and Jan Vondrak. On the hardness of welfare maximization in combinatorial auctions with submodular valuations. *SODA*, 2013.
- [3] Uriel Feige, Vahab S Mirrokni, and Jan Vondrak. Maximizing non-monotone submodular functions. *SIAM Journal on Computing*, 40(4):1133–1153, 2011.
- [4] Moran Feldman. Maximizing symmetric submodular functions. *arXiv preprint arXiv:1409.5900*, 2014.
- [5] Moran Feldman, Joseph Seffi Naor, and Roy Schwartz. Nonmonotone submodular maximization via a structural continuous greedy algorithm. In *Automata, Languages and Programming*, pages 342–353. Springer, 2011.
- [6] Shayan Oveis Gharan and Jan Vondrák. Submodular maximization by simulated annealing. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 1098–1116. SIAM, 2011.
- [7] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.
- [8] Benny Lehmann, Daniel Lehmann, and Noam Nisan. Combinatorial auctions with decreasing marginal utilities. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 18–28. ACM, 2001.
- [9] George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. An analysis of approximations for maximizing submodular set functions—i. *Mathematical Programming*, 14(1):265–294, 1978.